

国立国語研究所学術情報リポジトリ

[フロッピー版] 分類語彙表 国立国語研究所編 解説書

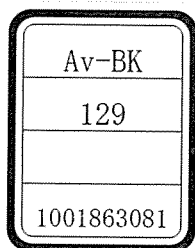
著者	国立国語研究所
ページ	1-6
発行年	1994-02-01
シリーズ	国立国語研究所言語処理データ集 ; 5
URL	http://doi.org/10.15084/00002342

国立国語研究所言語処理データ集 5

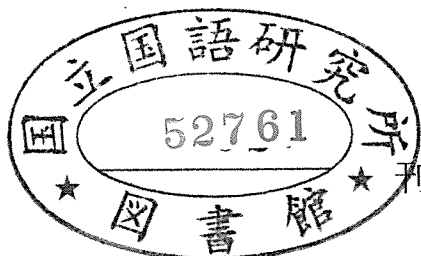
[フロッピー版]
分類語彙表

国立国語研究所 編

解説書



秀英出版



★ 刊行のことは

国立国語研究所言語処理データ集5として、『分類語彙表』[フロッピー版]を刊行します。

先に刊行した国立国語研究所資料集6『分類語彙表』は、収録語約3万2千6百を意味によって分類排列したのですが、本データ集は、これをほぼそのままフロッピーにうつしコンピュータによって利用できるようにしたものです。

『分類語彙表』は、現在29版をかさね、表現辞典としてはもとより広い分野での言語研究に利用されています。たとえば、基本語彙の選定のための基礎資料としての利用、ある言語作品の表現上の特色を見る物差しとしての利用、方言の分布の研究での利用、国語教育での利用などです。

さらに、最近ではこれらの研究にコンピュータを用いることが多く、また日本語情報処理の研究も盛んで、『分類語彙表』のフロッピー化を求める声がおおきくなっています。

現在、『分類語彙表』の改編を進めているところですが、これらの要求にこたえて刊行中のものと同じ内容のものを『分類語彙表』[フロッピー版]として刊行することとしました。

本データ集の作成は、言語体系研究部第2研究室の協力のもとに主として中野洋（前言語体系研究部第2研究室長、現言語体系研究部長）が担当し、小沼悦（言語体系研究部第2研究室研究補助員）がこれを助けました。なお、この研究および作業には文部省科学研究費「日本語シソーラス作成法の研究」（研究代表者：宮島達夫、昭和56,57年）、「言語データの収集と処理の研究」（研究代表者：野村雅昭、昭和61～63年）、「言語研究におけるシソーラスの利用法」（研究代表者：宮島達夫、平成1,2年）、「日本語シソーラス・データベースの構築とその利用法に関する研究」（研究代表者：中野洋、平成4,5年）の研究成果の一部を用いています。

平成5年12月20日

国立国語研究所

国立国語研究所長 水谷 修



1001863081

『分類語彙表』[フロッピー版] 解説書

1. データ公開の趣旨

国立国語研究所資料集6『分類語彙表』は、昭和39年3月に刊行されているが現在29版をかさねる。国立国語研究所の刊行物の中ではもっとも発行部数が多い。一般の表現辞典としての利用が多いためだろうが、言語研究への利用も少なくない。宮島達夫・小沼悦は「言語研究におけるシソーラスの利用法」(国立国語研究所報告104, 平成4年3月)で『分類語彙表』を言語研究に利用した論文119例を集めて解説している。そこに掲載されなかった論文の他、直接研究の対象や手段にはならなかったが、参考、目安として使われた研究など、『分類語彙表』を直接間接に利用した研究はこの何倍、何十倍にのぼると思われる。

国立国語研究所では昭和41年に電子計算機を導入し、各種の語彙調査に用いてきた。『高校教科書の語彙調査』(国立国語研究所報告87, 昭和61年3月)以降の語彙調査では分類語彙表の番号を語彙表に入れ、分析に用いた。『分類語彙表』の「まえがき」に予言されているとおり、最近、言語情報処理の研究では大規模なシソーラスが必要となり、『分類語彙表』を用いた研究が多くなっている。また、文科系の研究においてもコンピュータを用いることが多くなった。これを反映して、国立国語研究所にフロッピー版の『分類語彙表』の公開を求める声も大きくなった。

現在、『分類語彙表』の改編を進めているところであるが、当面の要求にこたえて刊行中のものと同じ内容のものを『分類語彙表』[フロッピー版]として刊行することとした。

2. 公開ファイル

公開するものは、フロッピー2枚である。

本解説書は、下記の解説ファイルを印刷したものにほぼ同じである。フロ

ッピーは2枚で以下のファイルを収める。

1 枚目	ファイル名
①フロッピー版解説ファイル	KAISETU.DOC
②まえがきファイル	MAEGAKI.TXT
③分類項目一覧ファイル	KOUMOKU.TXT
④分類語彙表ファイル	BUNRUI.DAT
2 枚目	
⑤五十音順索引ファイル	SAKUIN.DAT

それぞれMS-DOSのテキストファイルである。以下に詳細を示す。

3. まえがきファイル, 分類項目一覧ファイル

第29版そのままである。ただし、フランス語、ドイツ語文字でシフトJISコードにない記号は無視した。原本を参照していただきたい。

4. 分類語彙表ファイル

『分類語彙表』（以下表と略称する）をできるだけ忠実に再現した。ただし、以下の点は異なる。また、例として左に刊行中の『分類語彙表』を、右に『分類語彙表』〔フロッピー版〕を示す。

4.1 表の段落分けを数字で示した。また、1行の長さを45バイト未満とし、語は行をまたがないこととした。次に示す通りである。

[原 本]

[フロッピー版]

1.1635 朝 晩

朝晩 朝夕(161a) 昼夜 *午前 *午後・後
(こ) *正午
*暁 払暁 早暁 未明 明け 明け方 夜
明け前 *夜明け 有明 あけぼの しの
のめ 黎明 早朝 朝まだき 朝っぱら
昼まえ *朝
*昼間(ひるま) 昼間(ちゅうかん) 真昼間 昼
中 日中(にちゅう・ひなか) *白昼 日盛り
*昼 真昼 昼時 昼から 昼過ぎ 昼下
がり

1.1635 朝 晩

1 朝晩 朝夕 昼夜 *午前 *午後・後(こ) *正午
2 *暁 払暁 早暁 未明 明け 明け方 夜明け前
*夜明け 有明 あけぼの しののめ 黎明 早朝
朝まだき 朝っぱら 昼まえ *朝
3 昼間(ひるま) 昼間(ちゅうかん) 真昼間
昼中 日中(にちゅう・ひなか) *白昼
日盛り *昼 真昼 昼時 昼から 昼過ぎ
昼下がり

*暮れ 暮れ方 日暮れ 日の暮れ 日暮れ方
 方 *夕暮れ 夕景 日の入り 夕 *夕方
 夕刻 *タベ 夕まぐれ たそがれ 薄暮
 晩方 *晩 暮夜
 よい よいの口 春宵 *夜(よ・よる・や) 夜分
 *夜間 *夜中(やちゅう・よなか) 真夜中
 夜半 よわ 夜ふけ さ夜ふけ *深夜
 月夜 星月夜 おぼろ月夜 やみ夜 暗夜
 初更 三更 丑満時 初夜 中夜 後夜
 終日 夜通し 終夜 ひねもす よもすがら
 夜っびて 徹宵 宵越し
 ラッシュアワー

4 *暮れ 暮れ方 日暮れ 日の暮れ 日暮れ方
 *夕暮れ 夕景 日の入り 夕 *夕方 夕刻 *タベ
 夕まぐれ たそがれ 薄暮 晩方 *晩 暮夜
 5 よい よいの口 春宵 *夜(よ・よる・や) 夜分
 *夜間 *夜中(やちゅう・よなか) 真夜中 夜半
 よわ 夜ふけ さ夜ふけ *深夜 月夜 星月夜
 おぼろ月夜 やみ夜 暗夜 初更 三更 丑満時
 初夜 中夜 後夜
 6 終日 夜通し 終夜 ひねもす よもすがら
 夜っびて 徹宵 宵越し
 7 ラッシュアワー

4.2 2つ以上の分類番号で分けられる語群が、1つの分類項目の中に合併されている場合がある。この場合、〔フロッピー版〕では別々の分類番号の下に語を示し、分類項目名はつけない。

4.3 使用する文字は、分類番号、段落番号、段落内番号、付加情報注記のための括弧以外は全角文字を使用した。

また、次の文字は入力できないのでゲタ記号（シフト J I S の16進表示 81 A C）で表した。

罫 (1.183 の「罫罫」に用いられる)
 𪛗 (1.3121 の「𪛗𪛗」)
 𪛗 (1.551 の「𪛗葉」)
 𪛗 (2.3394 の「𪛗む」)
 𪛗 (3.5039 の「𪛗𪛗」)

4.4 語の表記は、計算機使用の便から次のように変えた。

原本では、「動詞は連用形を代表の形とし、五段活用は〈思いう〉のように、代表形のあとに連体形の語尾を小文字で添え、一段活用は〈起きル〉のように、代表形のあとにかたかな〈ル〉を小文字で添え、力変・サ変は、〈来くル〉〈関しル〉のように、代表形のあとに連体形を小文字で添え」である。これをフロッピー版では、下記右のように終止形で示した。次に、例を示す。

[原 本]

[フロッピー版]

2.111 関係

2.111 関係

*対しナル *属しナル 帰しナル *もとづき
く *従い 因りる *抛りる ちなみむ

1 *対する *属する 帰する *もとづく *従う

- 4.5 原本では「形容詞は語幹を代表形とし、〈赤い〉のように連体形語尾〈い〉を小文字で添えた。〈暖か〉のように、形容詞活用と形容動詞活用を持つものについては、〈暖かい. な〉のように、両方の連体形語尾〈い. な〉を小文字で添え」である。

フロッピー版では下記右のように、小文字で示された相の類の「い」は区別せず大文字で示した。形容詞活用と形容動詞活用を持つものは、形容詞活用の終止形「い」を大文字で、形容動詞活用の連体形語尾「な」は省略して示した。次に、例を示す。

[原 本]

[フロッピー版]

3.515 気象

3.515 気象

*熱い、はやはや *暑い、蒸し暑い、極暑
*暖かいな ほかほか ぽかぽか 生暖かい、

1 *熱い はやはや *暑い 蒸し暑い 極暑 *暖か
*暖かい ほかほか ぽかぽか 生暖かい

- 4.6 原本では下記左のように、「副詞的語句について、語尾の〈て〉〈と〉〈に〉等を小文字にしたものがある」が、フロッピー版では下記右のように、副詞的語句の語尾の「て」「と」「に」等はそれぞれ大文字で示した。次に、例を示す。

[原 本]

[フロッピー版]

3.15 変化・動き

3.15 変化・動き

おのずから おのずと ひとりでに 自然

1 おのずから おのずと ひとりでに 自然

5. 五十音順索引ファイル

五十音順索引ファイルは、以下に示すように項目やレコードを追加しており、『分類語彙表』の五十音順索引と異なる。レコード総数は36,780である。

5.1 五十音順索引ファイルのデータフォーマット

レコードは、次の7つの項目で構成されている。それぞれは半角のコンマ(,)で区切ってある。なお、ここに示す語の読み、表記は検索等のために設けたものであって、いずれもそれらの標準を示すものではない。

読み：ひらがなで示した。

表記：表での表記。

分類番号：表の通り。

段落番号：表で示された段落を数字で示した。

段落内番号：段落での語の順序を数字で示した。

付加情報：*, Y, Kの3種類の情報である。*は、それが表で*付きの語であることを示す。KまたはYがある語は、表には示されていない。Yは注記の語と異読であり、Kは注記の語と異表記であることを示す。

注記：表で表記とともに()付き等で示す情報。ただし、先頭が・の場合は、表で()なしで示される。

例を次に示す。

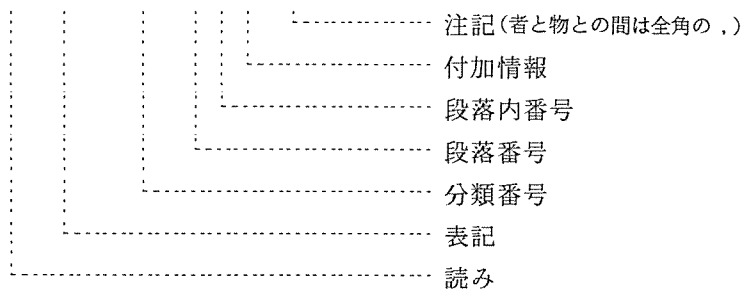
これ, これ, 1. 100, 1, 20, *,

の, の, 1. 100, 3, 30,, 私のです

きゃくたい, 客体, 1. 104, 1, 50,,

かくたい, 客体, 1. 104, 1, 60, Y, きゃくたい

もの, もの, 1. 100, 3, 10, *, 者, 物



5.2 レコードの追加

- ①「(お)じいさん」などは、この表記のレコードを作成し、さらにこの表記から想定される「おじいさん」のレコードも作った。
- ② 音便などで読みの異なるもの、同表記同義で異読のものなど五百余語を加えた。例えば、「秋風」は、索引に「あきかぜ」はあるが「しゅうふう」はないのでこれを加えた。

6. データ使用上の注意

本データは、国立国語研究所が保有する言語研究用データを電子計算機が読める形で公開する最初の刊行物である。これが言語研究の発展に寄与することを希望するとともに、この種の刊行を続けるために、以下の点についてご協力いただきたい。

- ①本データの利用に当たっては、国立国語研究所に著作権・版權があることを認め、その権利を侵害する行為は行わないこと。
- ②本データ作成のために各種の点検を何度も繰り返した。エラーを完全になくすには、なお点検が必要となろう。しかし、それでは公開が遅くなる。大量データの公開にはこの種の問題がつきまとう。そこで、利用者はこのことを承知の上でご利用願いたい。また、もしエラー等を発見された場合には、研究所にお知らせいただきたい。
- ③本データの利用による成果を公表する場合は、国立国語研究所の作成にかかるものであることを明示していただきたい。
- ④現在『分類語彙表』の改編作業が進行中である。